

AD-A086 192

MARYLAND UNIV COLLEGE PARK COMPUTER SCIENCE CENTER

F/G 12/1

THE BEHAVIOR OF A MULTIPLICITY INDEPENDENT ROOT-FINDING SCHEME --ETC(U)

APR 80 G W STEWART

N00014-76-C-0391

UNCLASSIFIED

TR-890

1 of 1
20-1
20-1

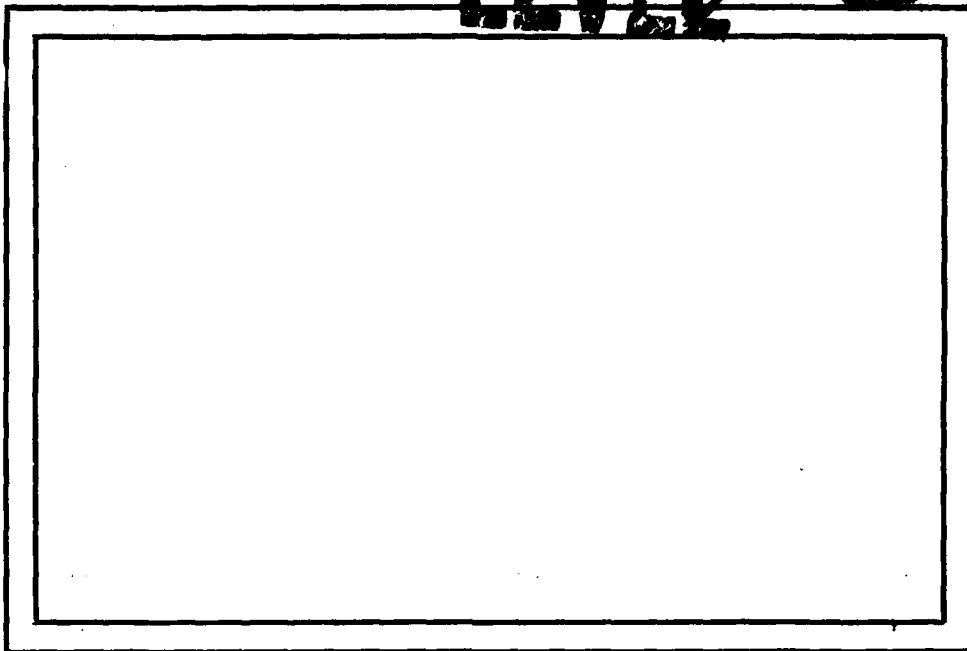
END	DATE FILMED	8-80	DTIC
-----	----------------	------	------

52

ADA 086192

LEVEL II

5



DTIC
ELECTE
JUN 27 1980
D
C

DDC FILE COPY

UNIVERSITY OF MARYLAND
COMPUTER SCIENCE CENTER

COLLEGE PARK, MARYLAND
20742

This document has been approved
for public release and sale; its
distribution is unlimited.

80 6 5 016

(14) TH-890

5

(9)

Technical Report, 890

April 1980

(12) 82

(6)

The Behavior of a Multiplicity
Independent Root-finding
Scheme in the Presence of Error.

(10)

G. W. Stewart*

(11) Apr 80

IC
STE
JUN 27 1980
D

Abstract

King [1] and Kioustelidis [2] have proposed a derivative free scheme *has been proposed* that permits root finding methods, such as the secant method to preserve high order convergence when the root in question is multiple. In this note it is shown that the scheme can fail to achieve the maximum accuracy that is attainable at a fixed precision of computation.

(15)

NA0014-76-C-0391

This document has been approved
for public release and sale; its
distribution is unlimited.

*This work was supported in part by the Office of Naval Research
under contract No. N00014-76-C-0391.

403018

alt

The Behavior of a Multiplicity
Independent Root-finding
Scheme in the Presence of Error

G. W. Stewart

The problem considered in this note is that of finding a root of the equation

$$f(x) = 0,$$

where f is a real valued function of a real variable. For simplicity we shall assume that f has a sufficient number of continuous derivatives for the purposes at hand.

There are any number of iterations, such as Newton's method or the secant method, that converge to a zero x^* of f from sufficiently near starting points. If x^* is a simple zero of f , i.e. if $f'(x^*) \neq 0$, then many of these iterations converge superlinearly; however, if x^* is a multiple zero, the convergence is usually linear and can be quite slow. One way of circumventing this difficulty, an approach which goes back at least to Schröder [3], is to replace f by a function g that is constructed from f in such a way that x^* is a simple zero of g . A classical choice is $g = f/f'$. Recently King [1] and Kioustelidis [2] have proposed the function

Accession For	
NTIS GMAI	
DDC TAB	
Unannounced	
Justification <i>per FL-82</i>	
By <i>file</i>	
Distribution/	
Availability Codes	
Dist	Availand/or special
<i>R</i>	

$$(1) \quad g(x) = \frac{[f(x)]^2}{f[x + f(x)] - f(x)},$$

where g and its derivatives are defined at x^* by continuity. It is easy to show that if x^* is a zero of multiplicity p of f then $g(x^*) = 0$ and

$$g'(x^*) = \frac{1}{p},$$

so that x^* is indeed a simple zero of g .

This technique is appealing when it is used with a method like the secant method, which does not require derivatives; at the cost of one extra function evaluation per iteration it promises rapid convergence, free from considerations of multiplicity. However, it is important to realize that the technique has a hidden cost. When f is evaluated with error, an iteration based on it will not get as near the zero as a more conventional method. It is the purpose of this note to show why this is so.

The general situation is adequately illustrated by the function

$$f(x) = cx^p,$$

which has a zero of multiplicity p at zero. Suppose that in practice we cannot evaluate $f(x)$ exactly but instead must work with a perturbed value

$$\tilde{f}(x) = f(x) + e(x),$$

where the error function $e(x)$ satisfies

$$|e(x)| \leq \epsilon.$$

Now in the interval

$$(2) \quad \left[-\left[\frac{\epsilon}{c}\right]^{\frac{1}{p}}, \left[\frac{\epsilon}{c}\right]^{\frac{1}{p}} \right]$$

the value of $|f(x)|$ is not greater than ϵ . Hence the value of \tilde{f} may be positive or negative at any point in the interval; i.e. any point in (2) is a potential zero. Wilkinson [4] has shown that if $e(x)$ is noisy, as it will be when it is due to rounding error, then an iterative method like Newton's method or the secant method will converge until it gets into the interval (2), after which it will behave erratically.

The same considerations apply to the function g ; however, we must now consider not only how accurately f is evaluated but also how accurately the denominator in (1) is evaluated. Even if the first function evaluation is done without error, we have for this denominator

$$\begin{aligned} & \tilde{f}[x + f(x)] - f(x) \\ &= c[x + cx^p]^p - cx^p + e[x + f(x)] \\ &= c^2 x^{2p-1} \left[p + \frac{p(p-1)}{2} x^{p-1} + \dots \right] + e[x + f(x)] \end{aligned}$$

As x approaches zero, the true value of the denominator becomes effectively $c^2 x^{2p-1}$, and this must be larger than ϵ to insure that it is at all accurate. In other words one should not expect further convergence once the interval

$$(3) \quad \left[- \left[\frac{\epsilon}{c^2} \right]^{\frac{1}{2p-1}}, \quad \left[\frac{\epsilon}{c^2} \right]^{\frac{1}{2p-1}} \right]$$

has been entered. For fixed c and sufficiently small ϵ , the interval (3) is always larger than (2).

A numerical experiment, performed on a Texas Instruments Program-mable 59 calculator, will illustrate the above considerations. The function f was taken to be

$$f(x) = (x-1)^3 = x^3 - 3x^2 + 3x - 1,$$

which has a triple zero at one. Since the TI 59 carries thirteen decimal digits in its computations, if Horner's method is used to evaluate f near one, we may expect an error on the order of 10^{-13} . Taking $\epsilon = 10^{-13}$ and $c = 1$, we get for the interval (2) (appropriately translated)

$$(4) \quad 1 + [-0.000046, +0.000046] ,$$

and for (3)

$$(5) \quad 1 + [-0.0025, +0.0025] .$$

The first interval is significantly smaller than the second.

Table 1 shows the results of applying the secant method, starting from $x_1 = 1.4$ and $x_2 = 1.2$, first to f and then to g . The convergence for f is slow but steady; the thirtieth iterate is just outside the interval (4), and at that point f evaluates exactly to zero. On the other hand the fourth iterate with the function g is in the interval (5). The iteration could not be carried further because the denominator in (1) evaluated exactly to zero.

The conclusion is that for multiple zeros the g -iteration can fail to achieve the maximum accuracy that is attainable at a fixed precision of computation. On the other hand, until the g -iteration enters its region of ineffectiveness, its behavior is quite good. In the example, it required the f -iteration eighteen function evaluations to produce a solution as accurate as the g -iteration produced in six. Perhaps a reasonable compromise is to use the g -iteration until it breaks down and then, if further accuracy is required, shift to more refined techniques.

References.

1. Richard F. King, "A Secant Method for Multiple Roots," BIT 17 (1977) 321-328
2. John B. Kioustelidis, "A Derivative-Free Transformation Preserving the Order of Convergence of Iteration Methods in Case of Multiple Zeros," Numer Math 33 (1979) 385-389
3. E. Schroeder, "Über unendliche viele Algorithmen zur Auflösung der Gleichungen," Math. Ann. 2(1870) 312-365
4. J. H. Wilkinson, "Rounding Errors in Algebraic Processes," Prentice-Hall, Englewood Cliffs, New Jersey, 1963.

Table 1

f- and g-Iterations for

$$f(x) = x^3 - 3x^2 + 3x - 1$$

Iterate	f	g
1	1.4000000	1.4000000
2	1.2000000	1.2000000
3	1.1714286	0.9438925
4	1.1228346	1.0016196
5	1.0945533	---
.	.	
.	.	
.	.	
18	1.0024305	
19	1.0018347	
20	1.0013850	
.	.	
.	.	
.	.	
27	1.0001933	
28	1.0001485	
29	1.0001150	
30	1.0000478	

